

Evaluation

Adam Bodnar

CPSC 533C
Monday, April 5, 2004

Motivation

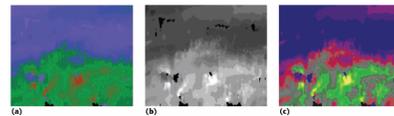
- So many techniques, so little evaluation
- Are they really effective?
 - How effective?
 - When are they effective?
 - Why are they effective?

Papers

- **User Studies: Why, How and When?** (Kosara et al., 2003)
- Navigation Patterns and Usability of Zoomable User Interfaces with and without an Overview (Hornbaek et al., 2002)
- An Evaluation of Information Visualization in Attention-Limited Environments (Somervell et al., 2002)

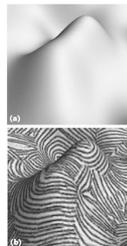
From Theory to Practice

- Can we design an effective colour sequence to illustrate features?
 - Chromatic sequence reveals categories (a)
 - Luminant sequence reveals form (b)



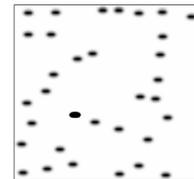
Comparison of Techniques

- Can we design an effective texture that conveys 3D shape information better than the current method?
 - Phong shading is default (a)
 - One principal direction texture mapping (b)



Study Within Context

- Can we effectively integrate semantic depth of field into an application?
 - Multi-layer map viewer
 - Layers can be opaque, semi-transparent, or SDOF
- No significant results



Other Techniques

- User studies aren't always the best choice
 - Time consuming, difficult to run, answer only small questions
- Field study
 - Observe the user in their native setting
- Visual designers
 - Replace part of user test with an expert

What to take away...

- Good experiments are difficult to design but are worth the effort
- User studies aren't always the most appropriate method of evaluation
- We need to establish evaluation as a standard InfoVis practice

Critique

- Strengths
 - Promotes evaluation through example
 - Accessible to those without a background in HCI
- Weaknesses
 - Only good points of studies presented
 - No critique of alternative evaluation techniques

Papers

- User Studies: Why, How and When? (Kosara et al., 2003)
- **Navigation Patterns and Usability of Zoomable User Interfaces with and without an Overview (Hornbaek et al., 2002)**
- An Evaluation of Information Visualization in Attention-Limited Environments (Somervell et al., 2002)

Experimental Background

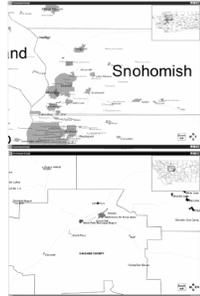
- Interfaces with an overview
 - Details of information space together with an overview of the entire information space
 - Established usability in literature
- Zoomable user interfaces
 - Organize information in space and scale, and use panning and zooming to navigate
 - Mixed results for usability in literature
- The usability of overviews for zoomable user interfaces **has not** been studied

What to Investigate?

- Question
 - How does the presence or absence of an overview in a zoomable interface affect usability?
- Hypotheses
 - Subjects will prefer the overview interface
 - The overview interface will be faster for comparison and browsing based tasks

Dataset and Tasks

- Dataset
 - Two maps based on census data
 - Differ in levels (single vs. multi-level)
- Tasks
 - Navigation and browsing



Study Design

- Experimental Design
 - Within 2 x 2 x 2 (interface, task, map)
 - Counterbalanced conditions
 - 32 subjects
- Measures
 - Quantitative
 - Accuracy, recall, speed, navigation actions
 - Qualitative
 - Preference, satisfaction

Results

- Significant Effects
 - Subjects preferred interface with an overview (**H1**)
 - Subjects faster with interface without an overview for multi-layer map (**H2**)
- Other
 - No difference between interfaces in subjects' ability to correctly solve tasks

Study Implications

- Consider the trade off between satisfaction and task completion time
- Unify overview with detail window
- Consider how map design influences usability

Critique

- Strengths
 - Detailed methodology
 - Real dataset and real test subjects
 - Strong statistical analysis and discussion
- Weaknesses
 - Investigators created the maps
 - No explanation for small display used in experiment

Papers

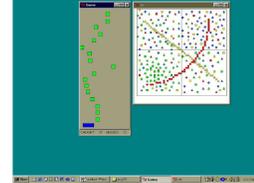
- User Studies: Why, How and When? (Kosara et al., 2003)
- Navigation Patterns and Usability of Zoomable User Interfaces with and without an Overview (Hornbaek et al., 2002)
- **An Evaluation of Information Visualization in Attention-Limited Environments (Somervell et al., 2002)**

What to Investigate?

- Motivation
 - InfoVis as a secondary display is a practical application but **has not** been evaluated
- Questions
 - How quickly and effectively can people interpret information visualization while busily performing other tasks?
 - What are the issues we must consider?

Experimental Setup

- Primary task
 - Video game
- Secondary task
 - Multiple choice questions about visualization target
 - Target could be single item or cluster



Study Design

- Experimental Design
 - Between/Within 2 x 2 x 2 (time, info density, task)
 - Counterbalanced conditions
 - 28 subjects
- Measures
 - Quantitative
 - Performance, correctness
 - Qualitative
 - None

Results

- Significant Effects
 - Subjects performed as good or better on low density visualizations vs. high density visualizations
 - Subjects achieved greater correctness (answering questions) when time = 8sec
- Other
 - No difference in primary task performance before or after the visualization appeared

Study Implications

- Peripheral visualizations can be introduced without hindering primary task performance
- Effective interpretation in a dual-task scenario requires more than one second
- Low information density displays result in performance that is as good as high density displays in a dual-task scenario

Critique

- Strengths
 - Ground experiment in previous work
 - Strong statistical analysis and discussion
- Weaknesses
 - Lack of real underlying data
 - Only focused on one type of primary task



Conclusion

- Empirical evaluation can lead to improvements in the design of information visualization
- Questions?